(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 March 2003 (20.03.2003)

PCT

(10) International Publication Number
WO 03/023010 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US02/28619

(22) International Filing Date: 6 September 2002 (06.09.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/317,879 7 September 2001 (07.09.2001) US

(71) Applicant: IRM LLC [US/US]; P. O. Box HM 2899, HM LX Hamilton (BM).

(72) Inventors: LIAO, Jiayu; 5032 Ashberry Road, Carlsbad, CA 92008 (US). GRAY, Nathanael, S.; 8674 Villa La Jolla Drive, Unit #3, San Diego, CA 92037 (US). CALDWELL, Jeremy, C.; 2031 Edinburg Avenue, Cardiff, CA 92007 (US). SCHULTZ, Peter, G.; 1650 La Jolla Rancho, San Diego, CA 92037 (US).

(74) Agents: HINSCH, Matthew, E. et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/023010 A2

(54) Title: SENSORY NEURON RECEPTORS

(57) Abstract: This invention provides G protein-coupled receptors that are expressed in mammalian sensory neurons of dorsal root ganglia. Both nucleic acids that encode the receptors, and the receptor polypeptides are provided.

Sensory Neuron Receptors

CROSS-REFERENCES TO RELATED APPLICATIONS

[01] The present application claims benefit of priority to U.S. Provisional Patent Application No. 60/317,879, filed September 7, 2001, which incorporated by reference.

BACKGROUND OF THE INVENTION

[02] The actions of many extracellular signals are mediated by the interaction of G-protein coupled receptors (GPCRs) and guanine nucleotide-binding regulatory proteins (G proteins). G protein-mediated signaling systems have been identified in many divergent organisms, such as mammals and yeast. GPCRs respond to, among other extracellular signals, neurotransmitters, hormones, odorants and light. GPCRs are similar and possess a number of highly conserved amino acids; the GPCRs are thought to represent a large "superfamily" of proteins. Individual GPCR types activate a particular signal transduction pathway; at least ten different signal transduction pathways are known to be activated via GPCRs.

[03] The cell bodies of sensory nerves that convey, *inter alia*, somatosensory (sense of touch) and nociceptive (sense of pain) information to the brain are found in the dorsal root ganglion. These ganglia (or "knotlike" masses) form swellings or nodules along the dorsal root at each level of the spinal cord. Dorsal root ganglia contain diverse subpopulations of primary sensory neurons. *See, e.g.*, Scott, S.A., SENSORY NEURONS: DIVERSITY, DEVELOPMENT AND PLASTICITY (Oxford University Press, 1992). The dorsal root ganglia contain both high- and low-threshold nociceptors.

[04] A number of nociceptive sensory receptors have not been identified. Identification of such receptors would be useful as target from screening for agents that regulate pain. The present invention addresses this and other problems.

BRIEF SUMMARY OF THE INVENTION

[05] The present invention provides isolated nucleic acids encoding a polypeptide at least 80% identical to SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12. In some embodiments, the polypeptide comprises

SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12. In some embodiments, the nucleic acids comprise SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9 or SEQ ID NO:11. In some embodiments, the nucleic acid encodes a receptor that has G-coupled protein receptor activity.

[06] The invention also provides isolated nucleic acids encoding an extracellular domain of a G-protein coupled receptor, the extracellular domain having at least 80% amino acid sequence identity to the extracellular domain of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12. In some embodiments, the extracellular domain comprises the extracellular domain of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

[07] The present invention also provides isolated nucleic acids encoding a transmembrane domain of a G-protein coupled receptor, the transmembrane domain comprising at least 80% amino acid sequence identity to the transmembrane domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12. In some embodiments, the transmembrane domain comprises the transmembrane domain of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

[08] The present invention also provides isolated polypeptides comprising a polypeptide sequence at least 80% identical to SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12. In some embodiments, the polypeptide is selected from SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12. In some embodiments, the receptor has G-protein coupled receptor activity.

[09] The present invention also provides isolated polypeptides comprising an extracellular domain of a G-protein coupled receptor, the extracellular domain comprising at least 80% amino acid sequence identity to the extracellular domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12. In some embodiments, the extracellular domain is the extracellular domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

[10] The present invention also provides isolated polypeptides comprising a transmembrane domain of a G-protein coupled receptor, the transmembrane domain comprising at least 80% amino acid sequence identity to the transmembrane domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12. In some embodiments, the extracellular domain is the extracellular domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

[11] The present invention also provides host cells comprising a heterologous polynucleotide encoding a polypeptide at least 80% identical to SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12. In some embodiments, the polypeptide comprises SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12. In some embodiments, the polynucleotide comprises SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9 or SEQ ID NO:11. In some embodiments, the host cell is selected from a COS cell, a CHO cell or a human embryonic kidney 293 cell. In some embodiments, the host cell is a eukaryotic cell. In some embodiments, the host cell is a prokaryotic cell.

[12] The present invention also provides antibodies that selectively bind to an isolated polypeptide selected from SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12..

[13] The present invention also provides an expression vector comprising a heterologous promoter operably linked to nucleic acids encoding a polypeptide at least 80% identical to SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

[14] The present invention also provides an expression vector comprising a heterologous promoter operably linked to nucleic acids encoding an extracellular domain of a G-protein coupled receptor, the extracellular domain having at least 80% amino acid sequence identity to the extracellular domain of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12

[15] The present invention also provides an expression vector comprising a heterologous promoter operably linked to the nucleic acid encoding a transmembrane domain of a G-protein coupled receptor, the transmembrane domain comprising at least 80% amino acid sequence identity to the transmembrane domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

[16] The present invention also provides a method for identifying a compound that modulates pain, the method comprising the steps of: (i) contacting a compound with a polypeptide comprising an extracellular domain of a G-protein coupled receptor, the extracellular domain comprising at least 80% amino acid sequence identity to the extracellular domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12; and (ii) selecting a compound that binds to the extracellular domain or that modulates G-protein coupled receptor activity of the polypeptide. In some embodiments, the method further comprises administering the selected compound to an

animal and determining the effect of the compound on pain sensitivity. In some embodiments, the polypeptide is a s G-protein coupled receptor, the receptor comprising at least 80% amino acid identity to SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12. In some embodiments, the polypeptide is SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

[17] In some embodiments, the polypeptide has G-protein coupled receptor activity. In some embodiments, the polypeptide is linked to a solid phase. In some embodiments, the polypeptide is expressed in a cell and the cell is contacted with the compound. In some embodiments, the G-protein coupled receptor activity of the polypeptide is determined by measuring changes in intracellular cAMP, IP3, or Ca^{2+}. In some embodiments, the method comprises selecting a compound that binds to the polypeptide.

DEFINITIONS

[18] As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

[19] A "Drg GPCR" refers to a G-protein coupled receptor, or fragment thereof, that is specifically expressed in dorsal root ganglia of a mammal (e.g., a human) and that are substantially identical to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12 or that are encoded by a polynucleotide that specifically hybridizes to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or SEQ ID NO:11. Optionally, the polypeptides of the invention are involved in sensory transduction, optionally pain transduction in dorsal root ganglia cells. The term encompasses polymorphic variants, alleles, mutants, and interspecies orthologs.

[20] "GPCR activity" refers to the ability of a GPCR to transduce a signal. Such activity can be measured in a heterologous cell, by coupling a GPCR (or a chimeric GPCR) to either a G-protein or promiscuous G-protein such as $G\alpha 15$, and an enzyme such as PLC, and measuring increases in intracellular calcium. *See, e.g.*, Offermans & Simon, *J. Biol. Chem.* 270:15175-15180 (1995). Receptor activity can be effectively measured by recording ligand-induced changes in Ca^{2+} concentration using fluorescent Ca^{2+}-indicator dyes and fluorometric imaging. G-protein coupled receptor activity includes a response of a G-protein coupled receptor (GPCR) to stimuli, e.g., the Drg GPCRs bind to G-proteins in response to extracellular stimuli and promote production of second messengers such as IP3, cAMP, and Ca^{2+} via stimulation of enzymes such as phospholipase C and adenylate cyclase.

Activity can be monitored by assaying, e.g., ligand binding, changes in ion flux, membrane potential, current flow, transcription, G-protein binding, GPCR phosphorylation or dephosphorylation, signal transduction, receptor-ligand interactions, second messenger concentrations (e.g., cAMP, IP3, or intracellular Ca^{2+}), *in vitro*, *in vivo*, and *ex vivo* and also includes other physiologic effects such increases or decreases of neurotransmitter or hormone release. Such functional effects can be measured by any means known to those skilled in the art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties, patch clamping, voltage-sensitive dyes, whole cell currents, radioisotope efflux, inducible markers, oocyte Drg GPCR expression; tissue culture cell Drg GPCR expression; transcriptional activation of Drg GPCR; ligand binding assays; voltage, membrane potential and conductance changes; ion flux assays; changes in intracellular second messengers such as cAMP and inositol triphosphate (IP3); changes in intracellular calcium levels; neurotransmitter release, and the like.

[21] Topologically, sensory GPCRs have an N-terminal "extracellular domain," a "transmembrane domain" comprising seven transmembrane regions and corresponding cytoplasmic and extracellular loops, and a C-terminal "cytoplasmic domain" (*see, e.g.*, Hoon *et al.*, *Cell* 96:541-551 (1999); Buck & Axel, *Cell* 65:175-187 (1991)). These domains can be structurally identified using methods known to those of skill in the art, such as sequence analysis programs that identify hydrophobic and hydrophilic domains (*see, e.g.*, Kyte & Doolittle, *J. Mol. Biol.* 157:105-132 (1982)). Such domains are useful for making chimeric proteins and for *in vitro* assays of the invention.

[22] "Extracellular domain" therefore refers to the domain of Drg GPCR that protrudes from the cellular membrane and binds to extracellular ligand. This region starts at the N-terminus and ends approximately at the leucine residue at amino acid position 30, 29, 29, 27, 29, and 27, for Drg1 (SEQ ID NO:2), Drg2 (SEQ ID NO:4), Drg4 (SEQ ID NO:6), Drg6 (SEQ ID NO:8), Drg7 (SEQ ID NO:10), and Drg8 (SEQ ID NO:12), respectively plus or minus approximately 5 amino acids.

[23] "Transmembrane domain," comprising seven transmembrane regions plus the corresponding cytoplasmic and extracellular loops, refers to the domain of Drg GPCR that starts approximately at the leucine residue at amino acid position 31, 30, 30, 28, 30, and 28 for Drg1, Drg2, Drg4, Drg6, Drg7,and Drg8, respectively plus or minus approximately 5 amino acids and ends approximately at the phenyalanine amino acid residue

at position 286, 286, 279, 280, 279, and 279 for Drg1, Drg2, Drg4, Drg6, Drg7,and Drg8, respectively plus or minus approximately 10 amino acids.

[24] "Cytoplasmic domain" refers to the domain of Drg GPCR that starts at the arginine at position 287, 287, 280, 281, 280, and 280 for Drg1, Drg2, Drg4, Drg6, Drg7,and Drg8, respectively plus or minus approximately 5 amino acids and continues to the C-terminus of the polypeptide.

[25] "Inhibitors," "activators," and "modulators" of Drg GPCR are used interchangeably to refer to inhibitory, activating, or modulating molecules identified using *in vitro* and *in vivo* assays for sensory (e.g., pain or somatosensory) transduction, e.g., ligands, agonists, antagonists, and their homologs and mimetics. The term "modulator" encompasses inhibitors and activators. Inhibitors are compounds that, e.g., bind to, partially or totally block stimulation, decrease, prevent, delay activation, inactivate, desensitize, or down regulate signal transduction, e.g., antagonists. Activators are compounds that, e.g., bind to, stimulate, increase, open, activate, facilitate, enhance activation, sensitize or up regulate signal transduction, e.g., agonists. Modulators include compounds that, e.g., alter the interaction of a receptor with: extracellular proteins that bind activators or inhibitor; G-proteins; kinases; and arrestin-like proteins, which also deactivate and desensitize receptors. Modulators include genetically modified versions of a Drg GPCR, e.g., with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, small chemical molecules and the like. Such assays for inhibitors and activators include, e.g., expressing Drg GPCR in cells or cell membranes, applying putative modulator compounds, and then determining the functional effects on sensory transduction. Samples or assays comprising a Drg GPCR that is treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative Drg GPCR activity value of 100%. Inhibition of Drg GPCR is achieved when the Drg GPCR activity value relative to the control is about 80%, optionally 50% or 25-0%. Activation of Drg GPCR is achieved when the Drg GPCR activity value relative to the control is 110%, optionally 150%, optionally 200-500%, or 1000-3000% higher.

[26] The terms "isolated" "purified" or "biologically pure" refer to material that is substantially or essentially free from components that normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein that is the predominant species present in a preparation is

substantially purified. The term "purified" denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the nucleic acid or protein is at least 85% pure, optionally at least 95% pure, and optionally at least 99% pure.

[27] "Nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs).

[28] Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* 260:2605-2608 (1985); Rossolini *et al.*, *Mol. Cell. Probes* 8:91-98 (1994)). The term "nucleic acid" encompasses the terms gene, cDNA, mRNA, oligonucleotide, and polynucleotide.

[29] The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

[30] The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, .gamma.-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an .alpha. carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified

peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

[31] "Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid that encodes a polypeptide is implicit in each described sequence.

[32] As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

[33] The following eight groups each contain amino acids that are conservative substitutions for one another:

1) Alanine (A), Glycine (G);

2) Aspartic acid (D), Glutamic acid (E);

3) Asparagine (N), Glutamine (Q);

4) Arginine (R), Lysine (K);

5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);

6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);

7) Serine (S), Threonine (T); and

8) Cysteine (C), Methionine (M) (*see, e.g.*, Creighton, *Proteins* (1984)).

[34] The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

[35] The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein indicates that the protein comprises two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

[36] A "promoter" is defined as an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

[37] An "expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression cassette can be part of

a plasmid, virus, or nucleic acid fragment. Typically, the expression cassette includes a nucleic acid to be transcribed operably linked to a promoter.

[38] The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same. Sequences are "substantially identical" of they have a specified percentage of amino acid residues or nucleotides that are the same (i.e., 70% identity, optionally 75%, 80%, 85%, 90%, or 95% identity over a specified region), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. This definition also encompasses the compliment of a test sequence. The present invention provides sequences substantially identical to, e.g., SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, or SEQ ID NO:12. Optionally, the identity exists over a region that is at least about 50 amino acids or nucleotides in length, or more preferably over a region that is 75-100 amino acids or nucleotides in length.

[39] For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

[40] A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), or by

manual alignment and visual inspection (see, e.g., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (Ausubel *et al.*, eds. 1995 supplement)).

[41] Two examples of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*; *Nuc. Acids Res.* 25:3389-3402 (1977) and Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990), respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al., supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always>0) and N penalty score for mismatching residues; always<0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) or 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

[42] The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in

a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001.

[43] An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequence.

[44] The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g., total cellular or library DNA or RNA).

[45] The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acid, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10° C. lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength pH. The Tm is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at Tm, 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (e.g., 10 to 50 nucleotides) and at least about 60° C. for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background,

optionally 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42° C, or, 5x SSC, 1% SDS, incubating at 65° C, with wash in 0.2x SSC, and 0.1% SDS at 65° C.

[46] Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37° C., and a wash in 1x SSC at 45° C. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency.

[47] "Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

[48] An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kDa) and one "heavy" chain (about 50-70 kDa). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

[49] Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce $F(ab)'_2$, a dimer of Fab which itself is a light chain joined to V.sub.H-C.sub.H1 by a disulfide bond. The $F(ab)'_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the $F(ab)'_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (see FUNDAMENTAL IMMUNOLOGY (Paul ed., 3d ed. 1993). While various antibody fragments are defined in terms

of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized de novo either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized de novo using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty *et al.*, *Nature* 348:552-554 (1990)).

[50] For preparation of monoclonal or polyclonal antibodies, any technique known in the art can be used (see, e.g., Kohler & Milstein, *Nature* 256:495-497 (1975); Kozbor *et al.*, *Immunology Today* 4:72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy* (1985)). Techniques for the production of single chain antibodies (U.S. Pat. No. 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (see, e.g., McCafferty *et al.*, *Nature* 348:552-554 (1990); Marks *et al.*, *Biotechnology* 10:779-783 (1992)).

[51] A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

[52] The phrase "specifically (or selectively) binds" to an antibody or "specifically (or selectively) immunoreactive with," when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein at least two times the background and do not substantially bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to Drg GPCR from specific species such as rat, mouse, or human can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with Drg GPCR and not with other proteins, except for polymorphic variants and alleles of Drg

GPCR. This selection may be achieved by subtracting out antibodies that cross-react with Drg GPCR molecules from other species. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (see, e.g., Harlow & Lane, Antibodies, A Laboratory Manual (1988), for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity). Typically a specific or selective reaction will be at least twice background signal or noise and more typically more than 10 to 100 times background.

[53] The phrase "selectively associates with" refers to the ability of a nucleic acid to "selectively hybridize" with another as defined above, or the ability of an antibody to "selectively (or specifically) bind to a protein, as defined above.

[54] By "host cell" is meant a cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa and the like, e.g., cultured cells, explants, and cells *in vivo*.

DETAILED DESCRIPTION OF THE INVENTION

I. Introduction

[55] The present invention provides nucleic acids encoding sensory neuron receptors specifically expressed in the dorsal root ganglia. These nucleic acids and the receptors that they encode are referred to as "Drg1-8 GPCRs" for "dorsal root ganglia G-protein coupled receptor." These dorsal root ganglia-specific GPCRs are components of the dorsal root ganglia sensory transduction system and mediate senses including, e.g., pain or somatosensory effects. These nucleic acids provide valuable probes for the identification of dorsal root ganglia cells, as the nucleic acids are specifically expressed in dorsal root ganglia.

[56] The invention also provides methods of screening for modulators, e.g., activators, inhibitors, stimulators, enhancers, agonists, and antagonists, of these Drg GPCRs. These modulators of sensory neurons are useful for pharmacological and genetic modulation of pain, heat, cold, touch and other sensory stimuli.

[57] The invention provides assays for sensory and pain modulation, where a Drg GPCR acts as an direct or indirect reporter molecule for the effect of sensory modulators. GPCRs can be used in assays, e.g., to measure changes in ion concentration, membrane potential, current flow, ion flux, transcription, signal transduction, receptor-ligand

interactions, second messenger concentrations, *in vitro*, *in vivo*, and *ex vivo*. In one embodiment, a Drg GPCR is used as an indirect reporter via attachment to a second reporter molecule such as green fluorescent protein (see, e.g., Mistili & Spector, *Nature Biotechnology* 15:961-964 (1997)). In another embodiment, a Drg GPCR is recombinantly expressed in a cell, and modulation of sensory transduction via GPCR activity is assayed by measuring changes in Ca^{2+} levels.

[58] Methods of assaying for modulators of sensory transduction include *in vitro* ligand binding assays using a Drg GPCR, portions thereof such as the extracellular domain, or chimeric proteins comprising one or more domains of a Drg GPCR, oocyte Drg GPCR expression; tissue culture cell Drg GPCR expression; transcriptional activation of a Drg GPCR; phosphorylation and dephosphorylation of GPCRs; G-protein binding to GPCRs; ligand binding assays; voltage, membrane potential and conductance changes; ion flux assays; changes in intracellular second messengers such as cAMP and inositol triphosphate; changes in intracellular calcium levels; and neurotransmitter release.

[59] The invention also provides for methods of detecting Drg GPCR nucleic acid and protein expression, allowing investigation of sensory (e.g., pain) transduction regulation.

[60] Specific regions of the Drg GPCR nucleotide and amino acid sequence may be used to identify polymorphic variants, interspecies homologs, and alleles of Drg GPCR. This identification can be made *in vitro*, e.g., under stringent hybridization conditions or PCR and sequencing, or by using the sequence information in a computer system for comparison with other nucleotide sequences. Typically, identification of polymorphic variants and alleles of Drg GPCR is made by comparing an amino acid sequence of about 25 amino acids or more, e.g., 50-100 amino acids. Amino acid identity of approximately at least 70% or above, optionally 80% or 90-95% or above typically demonstrates that a protein is a polymorphic variant, interspecies homolog, or allele of Drg GPCR. Sequence comparison can be performed using any of the sequence comparison algorithms discussed below. Antibodies that bind specifically to Drg GPCR or a conserved region thereof can also be used to identify alleles, interspecies homologs, and polymorphic variants.

[61] Drg GPCR nucleotide and amino acid sequence information may also be used to construct models of polypeptide interaction and regulation in a computer system. These models are subsequently used to identify compounds that can activate or inhibit Drg GPCR. Such compounds that modulate the activity of Drg GPCR can be used to investigate the role of Drg GPCR in sensory (e.g., pain) transduction.

[62] The isolation of Drg GPCR provides a means for assaying for inhibitors and activators of G-protein coupled receptor sensory transduction. Biologically active Drg GPCR is useful for testing inhibitors and activators of Drg GPCR as pain transducers or pain inhibitors using *in vivo* and *in vitro* expression that measure, e.g., transcriptional activation of Drg GPCR; ligand binding; phosphorylation and dephosphorylation; binding to G-proteins; G-protein activation; regulatory molecule binding; voltage, membrane potential and conductance changes; ion flux; intracellular second messengers such as cAMP and inositol triphosphate; intracellular calcium levels; and neurotransmitter release. Such activators and inhibitors identified using Drg GPCR, can be used to further study sensory transduction (e.g., nociception and somatosensory effects) and to identify specific pain agonists and antagonists.

II. Isolation of nucleic acids encoding Drg GPCR

[63] This invention relies on routine techniques in the field of recombinant genetics. Basic texts disclosing the general methods of use in this invention include Sambrook *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL (2nd ed. 1989); Kriegler, GENE TRANSFER AND EXPRESSION: A LABORATORY MANUAL (1990); and CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (Ausubel *et al.*, eds., 1994)).

[64] In general, the nucleic acid sequences encoding Drg GPCR and related nucleic acid sequence homologs are cloned from cDNA and genomic DNA libraries by hybridization with a probe, or isolated using amplification techniques with oligonucleotide primers. For example, Drg GPCR sequences are typically isolated from mammalian nucleic acid (genomic or cDNA) libraries by hybridizing with a nucleic acid probe, the sequence of which can be derived from SEQ ID NOs:1, 3, 5, 7, 9, or 11. A suitable tissue from which Drg GPCR RNA and cDNA can be isolated is dorsal root ganglia.

[65] Amplification techniques using primers can also be used to amplify and isolate Drg GPCR from DNA or RNA. Degenerate primers encoding the Drg GPCRs exemplified herein can also be used to amplify a sequence of Drg GPCR (*see, e.g.*, Dieffenfach & Dveksler, PCR PRIMER: A LABORATORY MANUAL (1995)). These primers can be used, e.g., to amplify either the full-length sequence or a probe of one to several hundred nucleotides, which is then used to screen a mammalian library for full-length Drg GPCR.

[66] Nucleic acids encoding Drg GPCR can also be isolated from expression libraries using antibodies as probes. Such polyclonal or monoclonal antibodies can be raised using the sequence of SEQ ID NOs: 2, 4, 6, 8, 10 or 12.

[67] Drg GPCR polymorphic variants, alleles, and interspecies homologs that are substantially identical to Drg GPCR can be isolated using Drg GPCR nucleic acid probes, and oligonucleotides under stringent hybridization conditions, by screening libraries. Alternatively, expression libraries can be used to clone Drg GPCR and Drg GPCR polymorphic variants, alleles, and interspecies homologs, by detecting expressed homologs immunologically with antisera or purified antibodies made against Drg GPCR, which also recognize and selectively bind to the Drg GPCR homolog.

[68] To make a cDNA library, one should choose a source that is rich in Drg GPCR mRNA, e.g., dorsal root ganglia. The mRNA is then made into cDNA using reverse transcriptase, ligated into a recombinant vector, and transfected into a recombinant host for propagation, screening and cloning. Methods for making and screening cDNA libraries are well known (see, e.g., Gubler & Hoffman, *Gene* 25:263-269 (1983); Sambrook *et al., supra*; Ausubel *et al., supra*).

[69] For a genomic library, the DNA is extracted from the tissue and either mechanically sheared or enzymatically digested to yield fragments of about 12-20 kb. The fragments are then separated by gradient centrifugation from undesired sizes and are constructed in bacteriophage lambda vectors. These vectors and phage are packaged *in vitro*. Recombinant phage are analyzed by plaque hybridization as described in Benton & Davis, *Science* 196:180-182 (1977). Colony hybridization is carried out as generally described in Grunstein *et al., Proc. Natl. Acad. Sci. USA*. 72:3961-3965 (1975).

[70] An alternative method of isolating Drg GPCR nucleic acid and its homologs combines the use of synthetic oligonucleotide primers and amplification of an RNA or DNA template (see U.S. Pat. Nos. 4,683,195 and 4,683,202; PCR PROTOCOLS: A GUIDE TO METHODS AND APPLICATIONS (Innis *et al.*, eds, 1990)). Methods such as polymerase chain reaction (PCR) and ligase chain reaction (LCR) can be used to amplify nucleic acid sequences of Drg GPCR directly from mRNA, from cDNA, from genomic libraries or cDNA libraries. Degenerate oligonucleotides can be designed to amplify Drg GPCR homologs using the sequences provided herein. Restriction endonuclease sites can be incorporated into the primers. Polymerase chain reaction or other in vitro amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of Drg GPCR encoding mRNA in physiological samples, for nucleic acid sequencing, or for other purposes. Genes amplified by the PCR reaction can be purified from agarose gels and cloned into an appropriate vector.

[71] Gene expression of Drg GPCR can also be analyzed by techniques known in the art, e.g., reverse transcription and amplification of mRNA, isolation of total RNA or poly A+ RNA, northern blotting, dot blotting, in situ hybridization, RNase protection, probing DNA microchip arrays, and the like. In one embodiment, high density oligonucleotide analysis technology (e.g., GeneChip™) is used to identify homologs and polymorphic variants of the GPCRs of the invention. In the case where the homologs being identified are linked to a known disease, they can be used with GeneChip™ as a diagnostic tool in detecting the disease in a biological sample. *See, e.g.*, Gunthand *et al.*, *AIDS Res. Hum. Retroviruses* 14: 869-876 (1998); Kozal *et al.*, *Nat. Med.* 2:753-759 (1996); Matson *et al.*, *Anal. Biochem.* 224:110-106 (1995); Lockhart *et al.*, *Nat. Biotechnol.* 14:1675-1680 (1996); Gingeras *et al.*, *Genome Res.* 8:435-448 (1998); Hacia *et al.*, *Nucleic Acids Res.* 26:3865-3866 (1998).

[72] Synthetic oligonucleotides can be used to construct recombinant Drg GPCR genes for use as probes or for expression of protein. This method is performed using a series of overlapping oligonucleotides usually 40-120 bp in length, representing both the sense and nonsense strands of the gene. These DNA fragments are then annealed, ligated and cloned. Alternatively, amplification techniques can be used with precise primers to amplify a specific subsequence of the Drg GPCR nucleic acid. The specific subsequence is then ligated into an expression vector.

[73] The nucleic acid encoding Drg GPCR is typically cloned into intermediate vectors before transformation into prokaryotic or eukaryotic cells for replication and/or expression. These intermediate vectors are typically prokaryote vectors, e.g., plasmids, or shuttle vectors.

[74] Optionally, nucleic acids encoding chimeric proteins comprising Drg GPCR or domains thereof can be made according to standard techniques. For example, a domain such as ligand binding domain, an extracellular domain, a transmembrane domain (e.g., one comprising seven transmembrane regions and corresponding extracellular and cytosolic loops), the transmembrane domain and a cytoplasmic domain, an active site, a subunit association region, etc., can be covalently linked to a heterologous protein. For example, an extracellular domain can be linked to a heterologous GPCR transmembrane domain, or a heterologous GPCR extracellular domain can be linked to a transmembrane domain. Other heterologous proteins of choice include, e.g., green fluorescent protein, β-gal, glutamate receptor, and the rhodopsin presequence.

[75] To obtain high level expression of a cloned gene or nucleic acid, such as those cDNAs encoding Drg GPCR, one typically subclones a Drg GPCR into an expression vector that contains a strong promoter to direct transcription, a transcription/translation terminator, and if for a nucleic acid encoding a protein, a ribosome binding site for translational initiation. Suitable bacterial promoters are well known in the art and described, e.g., in Sambrook *et al.* and Ausubel *et al.* Bacterial expression systems for expressing the Drg GPCR protein are available in, e.g., *E. coli*, *Bacillus* sp., and *Salmonella* (Palva *et al., Gene* 22:229-235 (1983); Mosbach *et al., Nature* 302:543-545 (1983). Kits for such expression systems are commercially available. Eukaryotic expression systems for mammalian cells, yeast, and insect cells are well known in the art and are also commercially available. In one embodiment, the eukaryotic expression vector is an adenoviral vector, an adeno-associated vector, or a retroviral vector.

[76] The promoter used to direct expression of a heterologous nucleic acid depends on the particular application. The promoter is optionally positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

[77] In addition to the promoter, the expression vector typically contains a transcription unit or expression cassette that contains all the additional elements required for the expression of the Drg GPCR encoding nucleic acid in host cells. A typical expression cassette thus contains a promoter operably linked to the nucleic acid sequence encoding Drg GPCR and signals required for efficient polyadenylation of the transcript, ribosome binding sites, and translation termination. The nucleic acid sequence encoding Drg GPCR may typically be linked to a cleavable signal peptide sequence to promote secretion of the encoded protein by the transformed cell. Such signal peptides would include, among others, the signal peptides from tissue plasminogen activator, insulin, and neuron growth factor, and juvenile hormone esterase of *Heliothis virescens*. Additional elements of the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns with functional splice donor and acceptor sites.

[78] In addition to a promoter sequence, the expression cassette can also contain a transcription termination region downstream of the structural gene to provide for efficient termination. The termination region may be obtained from the same gene as the promoter sequence or may be obtained from different genes.

[79] The particular expression vector used to transport the genetic information into the cell is not particularly critical. Any of the conventional vectors used for expression in eukaryotic or prokaryotic cells may be used. Standard bacterial expression vectors include plasmids such as pBR322 based plasmids, pSKF, pET23D, and fusion expression systems such as GST and LacZ. Epitope tags can also be added to recombinant proteins to provide convenient methods of isolation, e.g., c-myc.

[80] Expression vectors containing regulatory elements from eukaryotic viruses are typically used in eukaryotic expression vectors, e.g., SV40 vectors, papilloma virus vectors, and vectors derived from Epstein-Barr virus. Other exemplary eukaryotic vectors include pMSG, pAV009/A$^+$, pMTO10/A$^+$, pMAMneo-5, baculovirus pDSVE, and any other vector allowing expression of proteins under the direction of the SV40 early promoter, SV40 later promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, polyhedrin promoter, or other promoters shown effective for expression in eukaryotic cells.

[81] Some expression systems have markers that provide gene amplification such as thymidine kinase, hygromycin B phosphotransferase, and dihydrofolate reductase. Alternatively, high yield expression systems not involving gene amplification are also suitable, such as using a baculovirus vector in insect cells, with a Drg GPCR encoding sequence under the direction of the polyhedrin promoter or other strong baculovirus promoters.

[82] The elements that are typically included in expression vectors also include a replicon that functions in *E. coli*, a gene encoding antibiotic resistance to permit selection of bacteria that harbor recombinant plasmids, and unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. The particular antibiotic resistance gene chosen is not critical, any of the many resistance genes known in the art are suitable. The prokaryotic sequences are optionally chosen such that they do not interfere with the replication of the DNA in eukaryotic cells, if necessary.

[83] Standard transfection methods are used to produce bacterial, mammalian, yeast or insect cell lines that express large quantities of Drg GPCR protein, which are then purified using standard techniques (*see, e.g.*, Colley *et al.*, *J. Biol. Chem.* 264:17619-17622 (1989); *Guide to Protein Purification*, in METHODS IN ENZYMOLOGY, vol. 182 (Deutscher, ed., 1990)). Transformation of eukaryotic and prokaryotic cells are performed according to standard techniques (*see, e.g.*, Morrison, *J. Bact.* 132:349-351

(1977); Clark-Curtiss & Curtiss, Methods in Enzymology 101:347-362 (Wu *et al.*, eds, 1983).

[84] Any of the well known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, polybrene, protoplast fusion, electroporation, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (see, e.g., Sambrook *et al., supra*). It is only necessary that the particular genetic engineering procedure used be capable of successfully introducing at least one gene into the host cell capable of expressing Drg GPCR.

[85] After the expression vector is introduced into the cells, the transfected cells are cultured under conditions favoring expression of Drg GPCR, which is recovered from the culture using standard techniques identified below.

III. Purification of Drg GPCR

[86] Either naturally occurring or recombinant Drg GPCR can be purified for use in functional assays. Optionally, recombinant Drg GPCR is purified. Naturally occurring Drg GPCR is purified, e.g., from mammalian tissue such as dorsal root ganglia and any other source of a Drg GPCR homolog. Recombinant Drg GPCR is purified from any suitable bacterial or eukaryotic expression system, e.g., CHO cells or insect cells.

[87] Drg GPCR may be purified to substantial purity by standard techniques, including selective precipitation with such substances as ammonium sulfate; column chromatography, immunopurification methods, and others (see, e.g., Scopes, *Protein Purification: Principles and Practice* (1982); U.S. Pat. No. 4,673,641; Ausubel *et al., supra*; and Sambrook *et al., supra*).

[88] A number of procedures can be employed when recombinant Drg GPCR is being purified. For example, proteins having established molecular adhesion properties can be reversible fused to Drg GPCR. With the appropriate ligand, Drg GPCR can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form. The fused protein is then removed by enzymatic activity. Finally Drg GPCR could be purified using immunoaffinity columns.

A. Purification of Drg GPCRs from recombinant cells

[89] Recombinant proteins are expressed by transformed bacteria or eukaryotic cells such as CHO cells or insect cells in large amounts, typically after promoter induction; but expression can be constitutive. Promoter induction with IPTG is a one example of an inducible promoter system. Cells are grown according to standard procedures in the art. Fresh or frozen cells are used for isolation of protein.

[90] Proteins expressed in bacteria may form insoluble aggregates ("inclusion bodies"). Several protocols are suitable for purification of Drg GPCR inclusion bodies. For example, purification of inclusion bodies typically involves the extraction, separation and/or purification of inclusion bodies by disruption of bacterial cells, e.g., by incubation in a buffer of 50 mM TRIS/HCL pH 7.5, 50 mM NaCl, 5 mM MgCl.sub.2, 1 mM DTT, 0.1 mM ATP, and 1 mM PMSF. The cell suspension can be lysed using 2-3 passages through a French Press, homogenized using a Polytron (Brinkman Instruments) or sonicated on ice. Alternate methods of lysing bacteria are apparent to those of skill in the art (see, e.g., Sambrook *et al.*, *supra*; Ausubel *et al.*, *supra*).

[91] If necessary, the inclusion bodies are solubilized, and the lysed cell suspension is typically centrifuged to remove unwanted insoluble matter. Proteins that formed the inclusion bodies may be renatured by dilution or dialysis with a compatible buffer. Suitable solvents include, but are not limited to urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to about 8 M). Some solvents which are capable of solubilizing aggregate-forming proteins, for example SDS (sodium dodecyl sulfate), 70% formic acid, are inappropriate for use in this procedure due to the possibility of irreversible denaturation of the proteins, accompanied by a lack of immunogenicity and/or activity. Although guanidine hydrochloride and similar agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of immunologically and/or biologically active protein. Other suitable buffers are known to those skilled in the art. Drg GPCR is separated from other bacterial proteins by standard separation techniques, e.g., with Ni-NTA agarose resin.

[92] Alternatively, it is possible to purify Drg GPCR from bacteria periplasm. After lysis of the bacteria, when Drg GPCR is exported into the periplasm of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to skill in the art. To isolate recombinant proteins from the periplasm, the bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a

buffer containing 20% sucrose. To lyse the cells, the bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM $MgSO_4$ and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of skill in the art.

B. Protein separation techniques for purifying Drg GPCR

[93] Often as an initial step, particularly if the protein mixture is complex, an initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest. The preferred salt is ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol includes adding saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This concentration will precipitate the most hydrophobic of proteins. The precipitate is then discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, either through dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

[94] The molecular weight of Drg GPCR can be used to isolated it from proteins of greater and lesser size using ultrafiltration through membranes of different pore size (for example, Amicon or Millipore membranes). As a first step, the protein mixture is ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of the protein of interest. The retentate of the ultrafiltration is then ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

[95] Drg GPCR can also be separated from other proteins on the basis of its size, net surface charge, hydrophobicity, and affinity for ligands. In addition, antibodies raised against proteins can be conjugated to column matrices and the proteins immunopurified. All of these methods are well known in the art. It will be apparent to one

of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (e.g., Pharmacia Biotech).

V. Immunological detection of Drg GPCR

[96] In addition to the detection of Drg GPCR genes and gene expression using nucleic acid hybridization technology, one can also use immunoassays to detect Drg GPCR, e.g., to identify cells expressing Drg GPCRs and variants of Drg GPCR. Immunoassays can be used to qualitatively or quantitatively analyze Drg GPCR. A general overview of the applicable technology can be found in Harlow & Lane, ANTIBODIES: A LABORATORY MANUAL (1988).

[97] Methods of producing polyclonal and monoclonal antibodies that react specifically with Drg GPCR are known to those of skill in the art (*see, e.g.*, Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *supra*; Goding, MONOCLONAL ANTIBODIES: PRINCIPLES AND PRACTICE (2d ed. 1986); and Kohler & Milstein, *Nature* 256:495-497 (1975) . Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors, as well as preparation of polyclonal and monoclonal antibodies by immunizing rabbits or mice (see, e.g., Huse *et al.*, *Science* 246:1275-1281 (1989); Ward *et al.*, *Nature* 341:544-546 (1989)).

[98] A number of Drg GPCR comprising immunogens may be used to produce antibodies specifically reactive with Drg GPCR. For example, recombinant Drg GPCR or an antigenic fragment thereof, is isolated as described herein. Recombinant protein can be expressed in eukaryotic or prokaryotic cells as described above, and purified as generally described above. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used as an immunogen. Naturally occurring protein may also be used either in pure or impure form. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the protein.

[99] Methods of production of polyclonal antibodies are known to those of skill in the art. An inbred strain of mice (e.g., BALB/C mice) or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization protocol. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to Drg GPCR. When appropriately

high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (see Harlow & Lane, *supra*).

[100] Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*see* Kohler & Milstein, *Eur. J. Immunol.* 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse *et al.*, *Science* 246:1275-1281 (1989).

[101] Monoclonal antibodies and polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Typically, polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against non-Drg GPCR proteins or even other related proteins from other organisms, using a competitive binding immunoassay. Specific polyclonal antisera and monoclonal antibodies will usually bind with a K_d of at least about 0.1 mM, more usually at least about 1 μM, optionally at least about 0.1 μM or better, and optionally 0.01 μM or better.

[102] Once Drg GPCR specific antibodies are available, Drg GPCR can be detected by a variety of immunoassay methods. For a review of immunological and immunoassay procedures, see BASIC AND CLINICAL IMMUNOLOGY (Stites & Terr eds., 7th ed. 1991). Moreover, the immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in ENZYME IMMUNOASSAY (Maggio, ed., 1980); and Harlow & Lane, *supra*.

[103] Drg GPCR can be detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.*, U.S. Pat. Nos. 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see also *Methods in Cell Biology: Antibodies in Cell Biology*, volume 37 (Asai, ed. 1993); *Basic and*

Clinical Immunology (Stites & Terr, eds., 7th ed. 1991). Immunological binding assays (or immunoassays) typically use an antibody that specifically binds to a protein or antigen of choice (in this case the Drg GPCR or antigenic subsequence thereof). The antibody (e.g., anti-Drg GPCR) may be produced by any of a number of means well known to those of skill in the art and as described above.

[104] Immunoassays also often use a labeling agent to specifically bind to and label the complex formed by the antibody and antigen. The labeling agent may itself be one of the moieties comprising the antibody/antigen complex. Thus, the labeling agent may be a labeled Drg GPCR polypeptide or a labeled anti-Drg GPCR antibody. Alternatively, the labeling agent may be a third moiety, such a secondary antibody, that specifically binds to the antibody/Drg GPCR complex (a secondary antibody is typically specific to antibodies of the species from which the first antibody is derived). Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G may also be used as the label agent. These proteins exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (see, e.g., Kronval *et al.*, *J. Immunol.* 111:1401-1406 (1973); Akerstrom *et al.*, *J Immunol.* 135:2589-2542 (1985)). The labeling agent can be modified with a detectable moiety, such as biotin, to which another molecule can specifically bind, such as streptavidin. A variety of detectable moieties are well known to those skilled in the art.

[105] Throughout the assays, incubation and/or washing steps may be required after each combination of reagents. Incubation steps can vary from about 5 seconds to several hours, optionally from about 5 minutes to about 24 hours. However, the incubation time will depend upon the assay format, antigen, volume of solution, concentrations, and the like. Usually, the assays will be carried out at ambient temperature, although they can be conducted over a range of temperatures, such as 10° C. to 40° C.

[106] Immunoassays for detecting Drg GPCR in samples may be either competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of antigen is directly measured. In one "sandwich" assay, for example, the anti-Drg GPCR antibodies can be bound directly to a solid substrate on which they are immobilized. These immobilized antibodies then capture Drg GPCR present in the test sample. Drg GPCR is thus immobilized is then bound by a labeling agent, such as a second Drg GPCR antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second or third antibody is typically modified with a detectable

moiety, such as biotin, to which another molecule specifically binds, e.g., streptavidin, to provide a detectable moiety.

[107] In competitive assays, the amount of Drg GPCR present in the sample is measured indirectly by measuring the amount of a known, added (exogenous) Drg GPCR displaced (competed away) from an anti-Drg GPCR antibody by the unknown Drg GPCR present in a sample. In one competitive assay, a known amount of Drg GPCR is added to a sample and the sample is then contacted with an antibody that specifically binds to Drg GPCR. The amount of exogenous Drg GPCR bound to the antibody is inversely proportional to the concentration of Drg GPCR present in the sample. In some embodiments, the antibody is immobilized on a solid substrate. The amount of Drg GPCR bound to the antibody may be determined either by measuring the amount of Drg GPCR present in a Drg GPCR/antibody complex, or alternatively by measuring the amount of remaining uncomplexed protein. The amount of Drg GPCR may be detected by providing a labeled Drg GPCR molecule.

[108] In other embodiments, the ability of one or more antibodies to compete with a ligand of a Drg GPCR is determined. These assays can comprise measuring the ability of an antibody to compete with a labeled ligand for binding to a Drg GPCR of the invention.

[109] One of skill in the art will appreciate that it is often desirable to minimize non-specific binding in immunoassays. Particularly, where the assay involves an antigen or antibody immobilized on a solid substrate it is desirable to minimize the amount of non-specific binding to the substrate. Means of reducing such non-specific binding are well known to those of skill in the art. Typically, this technique involves coating the substrate with a proteinaceous composition. In particular, protein compositions such as bovine serum albumin (BSA), nonfat powdered milk, and gelatin are widely used with powdered milk being most preferred.

[110] The particular label or detectable group used in the assay is not a critical aspect of the invention, as long as it does not significantly interfere with the specific binding of the antibody used in the assay. The detectable group can be any material having a detectable physical or chemical property. Such detectable labels have been well-developed in the field of immunoassays and, in general, most any label useful in such methods can be applied to the present invention. Thus, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include magnetic beads (e.g., DYNAEADS™), fluorescent dyes (e.g., fluorescein isothiocyanate, Texas red, rhodamine, and the like), radiolabels (e.g., ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P), enzymes (e.g., horse radish peroxidase, alkaline

phosphatase and others commonly used in an ELISA), and calorimetric labels such as colloidal gold or colored glass or plastic beads (e.g., polystyrene, polypropylene, latex, etc.).

[111] The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. As indicated above, a wide variety of labels may be used, with the choice of label depending on sensitivity required, ease of conjugation with the compound, stability requirements, available instrumentation, and disposal provisions.

[112] Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to another molecules (e.g., streptavidin) molecule, which is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. The ligands and their targets can be used in any suitable combination with antibodies that recognize Drg GPCR, or secondary antibodies that recognize anti-Drg GPCR.

[113] The molecules can also be conjugated directly to signal generating compounds, e.g., by conjugation with an enzyme or fluorophore. Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidotases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazined- iones, e.g., luminol. For a review of various labeling or signal producing systems that may be used, see U.S. Pat. No. 4,391,904.

[114] Means of detecting labels are well known to those of skill in the art. Thus, for example, where the label is a radioactive label, means for detection include a scintillation counter or photographic film as in autoradiography. Where the label is a fluorescent label, it may be detected by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence. The fluorescence may be detected visually, by means of photographic film, by the use of electronic detectors such as charge coupled devices (CCDs) or photomnltipliers and the like. Similarly, enzymatic labels may be detected by providing the appropriate substrates for the enzyme and detecting the resulting reaction product. Finally simple calorimetric labels may be detected simply by observing the color associated with the label. Thus, in various dipstick assays, conjugated gold often appears pink, while various conjugated beads appear the color of the bead.

[115] Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labeled and the presence of the target antibody is detected by simple visual inspection.

V. Assays for Modulators of Drg GPCR

A. Assays for Drg GPCR activity

[116] Drg GPCR and its alleles and polymorphic variants are G-protein coupled receptors that participate in sensory (e.g., pain or somatosensory) transduction and hypertension control. The activity of Drg GPCR polypeptides can be assessed using a variety of *in vitro* and *in vivo* assays to determine functional, chemical, and physical effects, e.g., measuring ligand binding (e.g., radioactive ligand binding), second messengers (e.g., cAMP, cGMP, IP_3, DAG, or Ca^{2+}), ion flux, phosphorylation levels, transcription levels, neurotransmitter levels, and the like. Furthermore, such assays can be used to test for inhibitors and activators of Drg GPCR. Modulators can also be genetically altered versions of Drg GPCR.

[117] The Drg GPCR of the assay will be selected from a polypeptide having a sequence of SEQ ID NOs: 2, 4, 6, 8, 10, or 12 or a conservatively modified variant thereof. Alternatively, the Drg GPCR of the assay will be derived from a eukaryote and substantially identical to SEQ ID NOs: 2, 4, 6, 8, 10, or 12. Generally, the amino acid sequence identity will be at least 70%, optionally at least 85%, optionally at least 90-95%. Optionally, the polypeptide of the assays will comprise a domain of Drg GPCR, such as an extracellular domain, transmembrane domain, cytoplasmic domain, ligand binding domain, subunit association domain, active site, and the like. Either Drg GPCR or a domain thereof can be covalently linked to a heterologous protein to create a chimeric protein used in the assays described herein. IN some embodiments, a leader sequence is linked to the N-terminus of the Drg GPCR. Exemplary leader sequences included, e.g., the Igκ-chain V-J2-C signal peptide and the rhodopsin leadr sequence (amino acids 1-39). Surface expression of the GPCRs can be monitored, e.g., by immunofluorescence microscopy.

[118] Modulators of Drg GPCR activity are tested using Drg GPCR polypeptides as described above, either recombinant or naturally occurring. The protein can be isolated, expressed in a cell, expressed in a membrane derived from a cell, expressed in

tissue or in an animal, either recombinant or naturally occurring. Modulation is tested using one of the *in vitro* or *in vivo* assays described herein or as known to those in the art. Transduction can also be examined *in vitro* with soluble or solid state reactions, using a chimeric molecule such as an extracellular domain of a receptor covalently linked to a heterologous signal transduction domain, or a heterologous extracellular domain covalently linked to the transmembrane and or cytoplasmic domain of a receptor. Furthermore, ligand-binding domains of the protein of interest can be used in vitro in soluble or solid state reactions to assay for ligand binding.

[119] Ligand binding to Drg GPCR, a domain, or chimeric protein can be tested in solution, in a bilayer membrane, attached to a solid phase, in a lipid monolayer, or in vesicles. Exemplary Drg GPCR ligands include, e.g., Dynorphin A, Dynorphin A amide, Dynorphin A (1-6), Dynorphin A (1-13), Dynorphin A (2-13), Dynorphin A (2-17), Met-Enk, Met-Enk-RF-amide, Met-Enk-Arg-Phe, Met-Enk-Gly-leu, [D-pGlu1, D-Phe2, D-Trp3,6]-LH-RH, g1-MSH amide, g2-MSH, [N-MePhe1, D-Pro4]-Morphiceptin (PLO17), ACTH (Human), Leu-Enk, Adrenomedullin (22-52), Adrenomedullin (26-52) (Human)(ADM antagonist), Agouti 1-40 Amide, Agouti Related Protein (87-132)-Amide, Alpha-MSH, Alpha-Neo-Endorphin, Amylin Amide, BAM(1-20), BAM(1-22), BAM(2-22), BAM(6-22), BAM(1-20), ANP (Atrial Natriuretic Peptide), Anti-Inflamatory Peptide 1, Anti-Inflamatory Peptide 2, β-endorphin, Benzylureido-Met-Leu-Phe, Beta-ANP, Beta-Endorphin, Beta-MSH, Big Endothelin-1, Big Gastrin-1, BNP (Brain Natriuretic Peptide-32), BNP-45 (Cardiac Natriuretic Peptide, Bombesin, BAM(8-25), BAM(8-20), FLRF, Calcitonin Gene Related Peptide, NPFF, Calcitonin, Calcitonin Gene Related Peptide (8-37), CART (55-102), CART (55-102)[Met(O)67, CART (61-102), CGRP (8-37), CGRP II, Cholecystokinin Octapeptide [CCK(26-33)], Cholecystokinin-33, CNP-22 (C-Type Natriuretic Peptide), Corticotropin Releasing Factor, Cortistatin-14, NPAF, SST, NPY, FMRFamide, OrpaninFQFMRF amide related peptide, YMRFamide, YLPLRFamide, YFMRFamide, LPLRFamide, dFMRFamide, W-Nle-R-F-amide, and ACEP. Binding of a modulator can be tested using, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index) hydrodynamic (e.g., shape), chromatographic, or solubility properties. Competitive assays involving competition of a small molecule with a ligand for Drg GPCR are also provided.

[120] Receptor-G-protein interactions can also be examined. For example, binding of the G-protein to the receptor or its release from the receptor can be examined. For example, in the absence of GTP, an activator will lead to the formation of a tight complex of

a G protein (all three subunits) with the receptor. This complex can be detected in a variety of ways, as noted above. In the presence of GTP, release of the alpha subunit of the G protein from the other two G protein subunits serves as a criterion of activation. Such an assay can be modified to search for inhibitors.

[121] An activated or inhibited G-protein will in turn alter the properties of target enzymes, channels, and other effector proteins. The classic examples are the activation of cGMP phosphodiesterase by transducin in the visual system, adenylate cyclase by the stimulatory G-protein, phospholipase C by Gq and other cognate G proteins, and modulation of diverse channels by Gi and other G proteins. Downstream consequences can also be examined such as generation of diacyl glycerol and IP3 by phospholipase C, and in turn, for calcium mobilization by IP3.

[122] Activated GPCR receptors become substrates for kinases that phosphorylate the C-terminal tail of the receptor (and possibly other sites as well). Thus, activators will promote the transfer of ^{32}P from gamma-labeled GTP to the receptor, which can be assayed with a scintillation counter. The phosphorylation of the C-terminal tail will promote the binding of arrestin-like proteins and will interfere with the binding of G-proteins. The kinase/arrestin pathway plays a key role in the desensitization of many GPCR receptors. For example, compounds that modulate the duration a sensory receptor stays active would be useful as a means of prolonging a desired sensation or cutting off an unpleasant one (e.g., to reduce pain). For a general review of GPCR signal transduction and methods of assaying signal transduction, see, e.g., *Methods in Enzymology*, vols. 237 and 238 (1994) and volume 96 (1983); Bourne *et al.*, *Nature* 10:349:117-27 (1991); Bourne *et al.*, *Nature* 348:125-32 (1990); Pitcher *et al.*, *Annu. Rev. Biochem.* 67:653-92 (1998).

[123] Samples or assays that are treated with a potential Drg GPCR inhibitor or activator are compared to control samples without the test compound, to examine the extent of modulation. Control samples (untreated with activators or inhibitors) are assigned a relative Drg GPCR activity value of 100. Inhibition of Drg GPCR is achieved when the Drg GPCR activity value relative to the control is about 90%, optionally 50%, optionally 25-0%. Activation of Drg GPCR is achieved when the Drg GPCR activity value relative to the control is 110%, optionally 150%, 200-500%, or 1000-2000%.

[124] Changes in ion flux may be assessed by determining changes in polarization (i.e., electrical potential) of the cell or membrane expressing Drg GPCR. One method to determine changes in cellular polarization is by measuring changes in current (thereby measuring changes in polarization) with voltage-clamp and patch-clamp techniques,

e.g., the "cell-attached" mode, the "inside-out" mode, and the "whole cell" mode (see, e.g., Ackerman et al., New Engl. J. Med. 336:1575-1595 (1997)). Whole cell currents are conveniently determined using the standard methodology (*see, e.g.*, Hamil *et al.*, Pflugers. *Archiv.* 391:85 (1981). Other known assays include: radiolabeled ion flux assays and fluorescence assays using voltage-sensitive dyes (*see, e.g.*, Vestergarrd-Bogind *et al.*, *J. Membrane Biol.* 88:67-75 (1988); Gonzales & Tsien, *Chem.Biol.* 4:269-277 (1997); Daniel *et al.*, *J. Pharmacol. Meth.* 25:185-193 (1991); Holevinsky *et al.*, *J. Membrane Biology* 137:59-70 (1994)). Generally, the compounds to be tested are present in the range from 1 pM to 100 mM.

[125] The effects of the test compounds upon the function of the polypeptides can be measured by examining any of the parameters described above. Any suitable physiological change that affects GPCR activity can be used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as transmitter release, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as Ca^{2+}, IP3 or cAMP.

[126] Preferred assays for G-protein coupled receptors include cells that are loaded with ion or voltage sensitive dyes to report receptor activity. Assays for determining activity of such receptors can also use known agonists and antagonists for other G-protein coupled receptors as negative or positive controls to assess activity of tested compounds. In assays for identifying modulatory compounds (e.g., agonists, antagonists), changes in the level of ions in the cytoplasm or membrane voltage will be monitored using an ion sensitive or membrane voltage fluorescent indicator, respectively. Among the ion-sensitive indicators and voltage probes that may be employed are those disclosed in the Molecular Probes 1997 Catalog. For G-protein coupled receptors, promiscuous G-proteins such as Gq/i5, Gα15 and Gα16 can be used in the assay of choice (Wilkie *et al.*, *Proc. Nat'l Acad. Sci. USA* 88:10049-10053 (1991)). Such promiscuous G-proteins allow coupling of a wide range of receptors.

[127] Receptor activation typically initiates subsequent intracellular events, e.g., increases in second messengers such as IP3, which releases intracellular stores of calcium ions. Activation of some G-protein coupled receptors stimulates the formation of inositol triphosphate (IP3) through phospholipase C-mediated hydrolysis of

phosphatidylinositol (Berridge & Irvine, *Nature* 312:315-21 (1984)). IP3 in turn stimulates the release of intracellular calcium ion stores. Thus, a change in cytoplasmic calcium ion levels, or a change in second messenger levels such as IP3 can be used to assess G-protein coupled receptor function. Cells expressing such G-protein coupled receptors may exhibit increased cytoplasmic calcium levels as a result of contribution from both intracellular stores and via activation of ion channels, in which case it may be desirable although not necessary to conduct such assays in calcium-free buffer, optionally supplemented with a chelating agent such as EGTA, to distinguish fluorescence response resulting from calcium release from internal stores. Calcium stores can be measured using, e.g., the FLIPR system (Fluormetric Imaging Plate Reader, Molecular Devices Corp.).

[128] Other assays can involve determining the activity of receptors which, when activated, result in a change in the level of intracellular cyclic nucleotides, e.g., cAMP or cGMP, by activating or inhibiting enzymes such as adenylate cyclase. There are cyclic nucleotide-gated ion channels, e.g., rod photoreceptor cell channels and olfactory neuron channels that are permeable to cations upon activation by binding of cAMP or cGMP (see, e.g., Altenhofen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 88:9868-9872 (1991) and Dhallan *et al.*, *Nature* 347:184-187 (1990)). In cases where activation of the receptor results in a decrease in cyclic nucleotide levels, it may be preferable to expose the cells to agents that increase intracellular cyclic nucleotide levels, e.g., forskolin, prior to adding a receptor-activating compound to the cells in the assay. Cells for this type of assay can be made by co-transfection of a host cell with DNA encoding a cyclic nucleotide-crated ion channel, GPCR phosphatase and DNA encoding a receptor (e.g., certain glutamate receptors, muscarinic acetylcholine receptors, dopamine receptors, serotonin receptors, and the like), which, when activated, causes a change in cyclic nucleotide levels in the cytoplasm.

[129] In some embodiments, Drg GPCR activity is measured by expressing Drg GPCR in a heterologous cell with a promiscuous G-protein that links the receptor to a phospholipase C signal transduction pathway (see Offermans & Simon, J. Biol. Chem. 270:15175-15180 (1995)). Optionally the cell line is HEK-293 (which does not naturally express Drg GPCR) and the promiscuous G-protein is Gα15 (Offermans & Simon, *supra*). Modulation of sensory transduction is assayed by measuring changes in intracellular Ca^{2+} levels, which change in response to modulation of the Drg GPCR signal transduction pathway via administration of a molecule that associates with Drg GPCR. Changes in Ca^{2+}

levels are optionally measured using fluorescent Ca^{2+} indicator dyes and fluorometric imaging.

[130] In one embodiment, the changes in intracellular cAMP or cGMP can be measured using immunoassays. The method described in Offermans & Simon, *J. Biol. Chem.* 270:15175-15180 (1995) may be used to determine the level of cAMP. Also, the method described in Felley-Bosco *et al., Am. J. Resp. Cell and Mol. Biol.* 11:159-164 (1994) may be used to determine the level of cGMP. Further, an assay kit for measuring cAMP and/or cGMP is described in U.S. Pat. No. 4,115,538.

[131] In another embodiment, phosphatidyl inositol (PI) hydrolysis can be analyzed according to U.S. Pat. No. 5,436,128. Briefly, the assay involves labeling of cells with $_3$H-myoinositol for 48 or more hours. The labeled cells are treated with a test compound for one hour. The treated cells are lysed and extracted in chloroform-methanol-water after which the inositol phosphates were separated by ion exchange chromatography and quantified by scintillation counting. Fold stimulation is determined by calculating the ratio of cpm in the presence of agonist to cpm in the presence of buffer control. Likewise, fold inhibition is determined by calculating the ratio of cpm in the presence of antagonist to cpm in the presence of buffer control (which may or may not contain an agonist).

[132] In another embodiment, transcription levels can be measured to assess the effects of a test compound on signal transduction. A host cell containing the protein of interest is contacted with a test compound for a sufficient time to effect any interactions, and then the level of gene expression is measured. The amount of time to effect such interactions may be empirically determined, such as by running a time course and measuring the level of transcription as a function of time. The amount of transcription may be measured by using any method known to those of skill in the art to be suitable. For example, mRNA expression of the protein of interest may be detected using northern blots or their polypeptide products may be identified using immunoassays. Alternatively, transcription-based assays using reporter gene may be used as described in, e.g., U.S. Pat. No. 5,436,128. The reporter genes can be, e.g., chloramphenicol acetyltransferase, firefly luciferase, bacterial luciferase, β-galactosidase and alkaline phosphatase. Furthermore, the protein of interest can be used as an indirect reporter via attachment to a second reporter such as green fluorescent protein (see, e.g., Mistili & Spector, *Nature Biotechnology* 15:961-964 (1997)). Fluorescence can be monitored with, e.g., the Aquest system.

[133] The amount of transcription is then compared to the amount of transcription in either the same cell in the absence of the test compound, or it may be compared with the amount of transcription in a substantially identical cell that lacks the protein of interest. A substantially identical cell may be derived from the same cells from which the recombinant cell was prepared but which had not been modified by introduction of heterologous DNA. Any difference in the amount of transcription indicates that the test compound has in some manner altered the activity of the protein of interest.

[134] To further validate compounds identified in any of the methods described herein, lead candidates can be screen for analgesic or other desired qualities (e.g., hypertension control) in animals. In some embodiments, a lead candidate is tested in an animal model system for pain (e.g., as described in Malmberg *et al.*, *Science* 278: 279-283 (1997); Murata *et al.*, *Nature* 388:678-682 (1997)).

B. Modulators

[135] The compounds tested as modulators of Drg GPCR can be any small chemical compound, or a biological entity, such as a protein, sugar, nucleic acid or lipid. Alternatively, modulators can be genetically altered versions of Drg GPCR. Typically, test compounds will be small chemical molecules and peptides. Essentially any chemical compound can be used as a potential modulator or ligand in the assays of the invention, although most often compounds can be dissolved in aqueous or organic (especially DMSO-based) solutions are used. The assays are designed to screen large chemical libraries by automating the assay steps and providing compounds from any convenient source to assays, which are typically run in parallel (e.g., in microtiter formats on microtiter plates in robotic assays). It will be appreciated that there are many suppliers of chemical compounds, including Sigma (St. Louis, Mo.), Aldrich (St. Louis, Mo.), Sigma-Aldrich (St. Louis, Mo.), Fluka Chemika-Biochemica Analytika (Buchs Switzerland) and the like.

[136] In one embodiment, high throughput screening methods involve providing a combinatorial chemical or peptide library containing a large number of potential therapeutic compounds (potential modulator or ligand compounds). Such "combinatorial chemical libraries" or "ligand libraries" are then screened in one or more assays, as described herein, to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

[137] A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis, by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical building blocks (amino acids) in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

[138] Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Pat. No. 5,010,175, Furka, *Int. J. Pept. Prot. Res.* 37:487-493 (1991) and Houghton *et al., Nature* 354:84-88 (1991)). Other chemistries for generating chemical diversity libraries can also be used. Such chemistries include, but are not limited to: peptoids (e.g., PCT Publication No. WO 91/19735), encoded peptides (e.g., PCT Publication WO 93/20242), random bio-oligomers (e.g., PCT Publication No. WO 92/00091), benzodiazepines (e.g., U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al., Proc. Nat. Acad. Sci. USA* 90:6909-6913 (1993)), vinylogous polypeptides (Hagihara *et al, J. Amer. Chem. Soc.* 114:6568 (1992)), nonpeptidal peptidomimetics with glucose scaffolding (Hirschmann et al., *J. Amer. Chem.* Soc. 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen *et al., J. Amer. Chem. Soc.* 116:2661 (1994)), oligocarbamates (Cho *et al., Science* 261:1303 (1993)), and/or peptidyl phosphonates (Campbell *et al., J. Org. Chem.* 59:658 (1994)), nucleic acid libraries (see Ausubel, Berger and Sambrook, all *supra*), peptide nucleic acid libraries (see, e.g., U.S. Pat. No. 5,539,083), antibody libraries (see, e.g., Vaughn *et al., Nature Biotechnology*, 14(3):309-314 (1996) and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang *et al., Science*, 274:1520-1522 (1996) and U.S. Pat. No. 5,593,853), small organic molecule libraries (see, e.g., benzodiazepines, Baum C&EN, January 18, page 33 (1993); isoprenoids, U.S. Pat. No. 5,569,588; thiazolidinones and metathiazanones, U.S. Pat. No. 5,549,974; pyrrolidines, U.S. Pat. Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Pat. No. 5,506,337; benzodiazepines, U.S. Pat. No. 5,288,514, and the like).

[139] Devices for the preparation of combinatorial libraries are commercially available (see, e.g., 357 MPS, 390 MPS, Advanced Chem Tech, Louisville Ky., Symphony, Rainin, Woburn, Mass., 433A Applied Biosystems, Foster City, Calif., 9050 Plus, Millipore, Bedford, Mass.). In addition, numerous combinatorial libraries are themselves commercially

available (see, e.g., ComGenex, Princeton, N.J., Tripos, Inc., St. Louis, Mo., 3D Pharmaceuticals, Exton, Pa., Martek Biosciences, Columbia, Md., etc.).

C. Solid State and soluble high throughput assays

[140] Some embodiments the invention provide soluble assays using molecules such as a domain such as ligand binding domain, an extracellular domain, a transmembrane domain (e.g., one comprising seven transmembrane regions and cytosolic loops), the transmembrane domain and a cytoplasmic domain, an active site, a subunit association region, etc.; a domain that is covalently linked to a heterologous protein to create a chimeric molecule; a full-length Drg GPCR; or a cell or tissue expressing Drg GPCR, either naturally occurring or recombinant. In another embodiment, the invention provides solid phase based *in vitro* assays in a high throughput format, where the domain, chimeric molecule, Drg GPCR, or cell or tissue expressing Drg GPCR is attached to a solid phase substrate.

[141] In the high throughput assays of the invention, it is possible to screen up to several thousand different modulators or ligands in a single day. In particular, each well of a microtiter plate can be used to run a separate assay against a selected potential modulator, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 100 (e.g., 96) modulators. If 1536 well plates are used, then a single plate can easily assay from about 100- about 1500 different compounds. It is possible to assay several different plates per day; assay screens for up to about 6,000-20,000 different compounds is possible using the integrated systems of the invention. More recently, microfluidic approaches to reagent manipulation have been developed, e.g., by Caliper Technologies (Palo Alto, Calif.).

[142] The molecule of interest can be bound to the solid state component, directly or indirectly, via covalent or non covalent linkage, e.g., via a tag. The tag can be any of a variety of components. In general, a molecule which binds the tag (a tag binder) is fixed to a solid support, and the tagged molecule of interest (e.g., a Drg GPCR) is attached to the solid support by interaction of the tag and the tag binder.

[143] A number of tags and tag binders can be used, based upon known molecular interactions well described in the literature. For example, where a tag has a natural binder, for example, biotin, protein A, or protein G, it can be used in conjunction with appropriate tag binders (avidin, streptavidin, neutravidin, the Fc region of an immunoglobulin, etc.). Antibodies specific for molecules with natural binders such as biotin

are also widely available and appropriate tag binders. *See, e.g.*, SIGMA Immunochemicals 1998 catalogue SIGMA, St. Louis Mo.).

[144] Similarly, any haptenic or antigenic compound can be used in combination with an appropriate antibody to form a tag/tag binder pair. Thousands of specific antibodies are commercially available and many additional antibodies are described in the literature. For example, in one common configuration, the tag is a first antibody and the tag binder is a second antibody which recognizes the first antibody. In addition to antibody-antigen interactions, receptor-ligand interactions are also appropriate as tag and tag-binder pairs. For example, agonists and antagonists of cell membrane receptors (e.g., cell receptor-ligand interactions such as transferrin, c-kit, viral receptor ligands, cytokine receptors, chemokine receptors, interleukin receptors, immunoglobulin receptors and antibodies, the cadherein family, the integrin family, the selectin family, and the like; see, e.g., Pigott & Power, The Adhesion Molecule Facts Book I (1993)) can be used. Similarly, toxins and venoms, viral epitopes, hormones (e.g., opiates, steroids, etc.), intracellular receptors (e.g. which mediate the effects of various small ligands, including steroids, thyroid hormone, retinoids and vitamin D; peptides), drugs, lectins, sugars, nucleic acids (both linear and cyclic polymer configurations), oligosaccharides, proteins, phospholipids and antibodies can all interact with various cell receptors.

[145] Synthetic polymers, such as polyurethanes, polyesters, polycarbonates, polyureas, polyamides, polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides, and polyacetates can also form an appropriate tag or tag binder. Many other tag/tag binder pairs are also useful in assay systems described herein, as would be apparent to one of skill upon review of this disclosure.

[146] Common linkers such as peptides, polyethers, and the like can also serve as tags, and include polypeptide sequences, such as poly-gly sequences of between about 5 and 200 amino acids. Such flexible linkers are known to persons of skill in the art. For example, poly(ethelyne glycol) linkers are available from Shearwater Polymers, Inc. Huntsville, Alabama. These linkers optionally have amide linkages, sulfhydryl linkages, or heterofunctional linkages.

[147] Tag binders are fixed to solid substrates using any of a variety of methods currently available. Solid substrates are commonly derivatized or functionalized by exposing all or a portion of the substrate to a chemical reagent that fixes a chemical group to the surface that is reactive with a portion of the tag binder. For example, groups that are suitable for attachment to a longer chain portion would include amines, hydroxyl, thiol, and

carboxyl groups. Aminoalkylsilanes and hydroxyalkylsilanes can be used to functionalize a variety of surfaces, such as glass surfaces. The construction of such solid phase biopolymer arrays is well described in the literature. *See, e.g.*, Merrifield, *J. Am. Chem. Soc.* 85:2149-2154 (1963) (describing solid phase synthesis of, e.g., peptides); Geysen *et al.*, *J. Immun. Meth.* 102:259-274 (1987) (describing synthesis of solid phase components on pins); Frank & Doring, *Tetrahedron* 44:60316040 (1988) (describing synthesis of various peptide sequences on cellulose disks); Fodor *et al.*, *Science*, 251:767-777 (1991); Sheldon *et al.*, *Clinical Chemistry* 39(4):718-719 (1993); and Kozal *et al.*, *Nature Medicine* 2(7):753759-(1996) (all describing arrays of biopolymers fixed to solid substrates). Non-chemical approaches for fixing tag binders to substrates include other common methods, such as heat, cross-linking by UV radiation, and the like.

VI. Kits

[148] Drg GPCR and its homologs are a useful tool for identifying dorsal root ganglia cells, examining sensory transduction and identifying new analgesic compounds. Drg GPCR-specific reagents that specifically hybridize to Drg GPCR nucleic acid, such as Drg GPCR probes and primers, and Drg GPCR specific reagents that specifically bind to the Drg GPCR protein, e.g., Drg GPCR antibodies are used to examine sensory transduction regulation.

[149] Nucleic acid assays for the presence of Drg GPCR DNA and RNA in a sample include numerous techniques are known to those skilled in the art, such as Southern analysis, northern analysis, dot blots, RNase protection, S1 analysis, amplification techniques such as PCR and LCR, and *in situ* hybridization.

[150] The present invention also provides for kits for screening for modulators of a Drg GPCR. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise any one or more of the following materials: a Drg GPCR, reaction tubes, and instructions for testing Drg GPCR activity. Optionally, the kit contains biologically active Drg GPCR. A wide variety of kits and components can be prepared according to the present invention, depending upon the intended user of the kit and the particular needs of the user.

VII. Administration and Pharmaceutical Compositions

[151] Modulators of the invention can be administered directly to the mammalian subject for modulation of pain *in vivo*. Administration is by any of the routes

normally used for introducing a modulator compound into ultimate contact with the tissue to be treated. The modulators are administered in any suitable manner, optionally with pharmaceutically acceptable carriers. Suitable methods of administering such modulators are available and well known to those of skill in the art, and, although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route.

[152] Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition. Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions of the present invention (see, e.g., *Remington's Pharmaceutical Sciences*, 17th ed. 1985)).

[153] The modulators, alone or in combination with other suitable components, can be made into aerosol formulations (i.e., they can be "nebulized") to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

[154] Formulations suitable for administration include aqueous and non-aqueous solutions, isotonic sterile solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. In the practice of this invention, compositions can be administered, for example, by orally, topically, intra[illegible]sly, intraperitoneally, intravesically or intrathecally. Optionally, the compositions are administered nasally. The formulations of compounds can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials. Solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described. The m[illegible] can also be administered as part a of prepared food or drug. The compounds of the present invention can also be used effectively in combination with one or more additional active [illegible] (e.g., other analgesics or anesthetics) depending on the desired therapy or effect.

[155] The dose administered to a patient, in the context of the present invention should be sufficient to effect a benefi[illegible]se in the subject over time. The dose will be determined by the efficacy of the p[illegible] modulators employed and the condition of the subject, as well as the body we[illegible]face area of the area to be treated. The size of the dose also will be determined by [illegible]ce, nature, and extent of any

adverse side-effects that accompany the administration of a particular compound or vector in a particular subject.

[156] In determining the effective amount of the modulator to be administered in a physician may evaluate circulating plasma levels of the modulator, modulator toxicities, and the production of anti-modulator antibodies. In general, the dose equivalent of a modulator is from about 1 ng/kg to 10 mg/kg for a typical subject.

[157] For administration, modulators of the present invention can be administered at a rate determined by the LD-50 of the modulator, and the side-effects of the inhibitor at various concentrations, as applied to the mass and overall health of the subject. Administration can be accomplished via single or divided doses.

[158] All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

[159] Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to one of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

EXAMPLES

[160] The following examples are offered to illustrate, but not to limit the claimed invention.

[161] We examined over 50 human tissues for the expression patterns of Drg GPCR members. Drg 6 (SEQ ID NO:8) is uniquely expressed in dorsal root ganglion. All other members of Drg family GPCR are also expressed in dorsal root ganglion, but not solely restricted in dorsal root ganglion. For example, Drg 4 (SEQ ID NO:6)is also expressed in heart and uterus corpus, and Drg 1 (SEQ ID NO:2) is expressed in amygdale, salivary gland and pons, while Drg2 (SEQ ID NO:4) is more widely expressed in many tissues. All the Human Drg family of GPCRs have been cloned from various resources, including from genomic DNA.

[162] The human tissues used for examination of Drg GPCR expression included: whole brain, placenta, heart, thyroid, [illegible] uterus, leukemia, promyelocytic-HL-60, spinal cord, fetal liver, salivary gland, a[illegible], prostate, cerebellum, liver, lung,

trachea, thymus, kidney, bone marrow, pituitary, adrenal cortex, dorsal root ganglion, lymph node GALT, ovary, parathyroid, retina, tonsil, occipital lobe, cerebellum peduncles, parietal lobe, cingulate cortex, substantia innominate, substantia nigra, olfactory bulb, optic chiasma atrioventricular node, ciliary ganglion, globus pallidus, sinotrial node, skin, subthalamic nucleus, superior cervical ganglion, medulla oblongata, pons, bladder, appendix, trigeminal ganglion, temporal lobe, pineal gland, tongue, mammary gland, uterus corpus, and skeletal muscle.

[163] The nerves that send somatosensory (sense of touch) and nociceptive (sense of pain) information to the brain are found in the dorsal root ganglion. As the Drg GPCR are specifically expressed in dorsal root ganglia, modulators of the Dgr GPCRs will modulate pain and somatosensory reception and transduction.

[164] In the snail *Helix aspersa* and in *Aplysia*, FMRFamide induces a fast depolarizing response due to direct activation of an amiloride-sensitive Na^+ channel. FMRFamide has been demonstrated in the involvement of regulating the heartbeat of a pulmonate mollusk. The human RFamide-family of neuropeptides can also activate DRG family member(s) in heart and increase the blood pressure. Our data suggests that at least a member of DRG family GPCRs are expressed in the heart (DRG4). Therefore, the antagonists of such member are valuable for hypertension control.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid encoding a polypeptide at least 80% identical to SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

2. The isolated nucleic acid of claim 1, wherein the polypeptide comprises SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

3. The isolated nucleic acid of claim 1, comprising SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9 or SEQ ID NO:11.

4. The isolated nucleic acid of claim 1, wherein the nucleic acid encodes a receptor that has G-coupled protein receptor activity.

5. An isolated nucleic acid encoding an extracellular domain of a G-protein coupled receptor, the extracellular domain having at least 80% amino acid sequence identity to the extracellular domain of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

6. The isolated nucleic acid of claim 5, wherein the extracellular domain comprises the extracellular domain of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

7. An isolated nucleic acid encoding a transmembrane domain of a G-protein coupled receptor, the transmembrane domain comprising at least 80% amino acid sequence identity to the transmembrane domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

8. The isolated nucleic acid of claim 7, wherein the transmembrane domain comprises the transmembrane domain of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

9. An isolated polypeptide comprising a polypeptide sequence at least 80% identical to SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

10. The isolated polypeptide of claim 9, wherein the polypeptide is selected from SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

11. The isolated polypeptide of claim 9, wherein the receptor has G-protein coupled receptor activity.

12. An isolated polypeptide comprising an extracellular domain of a G-protein coupled receptor, the extracellular domain comprising at least 80% amino acid sequence identity to the extracellular domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

13. The isolated polypeptide of claim 12, wherein the extracellular domain is the extracellular domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

14. An isolated polypeptide comprising a transmembrane domain of a G-protein coupled receptor, the transmembrane domain comprising at least 80% amino acid sequence identity to the transmembrane domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

15. The isolated polypeptide of claim 14, wherein the extracellular domain is the extracellular domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

16. A host cell comprising a heterologous polynucleotide encoding a polypeptide at least 80% identical to SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

17. The host cell of claim 16, wherein the polypeptide comprises SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

18. The host cell of claim 16, wherein the polynucleotide comprises SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9 or SEQ ID NO:11.

19. The host cell of claim 16, wherein the host cell is selected from a COS cell, a CHO cell or a human embryonic kidney 293 cell.

20. The host cell of claim 16, wherein the host cell is a eukaryotic cell.

21. The host cell of claim 16, wherein the host cell is a prokaryotic cell.

22. An antibody that selectively binds to the polypeptide of claim 10.

23 An expression vector comprising the nucleic acid of claim 1.

24. An expression vector comprising the nucleic acid of claim 5.

25. An expression vector comprising the nucleic acid of claim 7.

26. A method for identifying a compound that modulates pain, the method comprising the steps of: (i) contacting a compound with a polypeptide comprising an extracellular domain of a G-protein coupled receptor, the extracellular domain comprising at least 80% amino acid sequence identity to the extracellular domain of SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12; and (ii) selecting a compound that binds to the extracellular domain or that modulates G-protein coupled receptor activity of the polypeptide.

27. The method of claim 26, further comprising administering the selected compound to an animal and determining the effect of the compound on pain sensitivity.

28. The method of claim 26, wherein the polypeptide is a s G-protein coupled receptor, the receptor comprising at least 80% amino acid identity to SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

29. The method of claim 28, wherein the polypeptide is SEQ ID NO:2, SEQ ID NO;4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:12.

30. The method of claim 26, wherein the polypeptide has G-protein coupled receptor activity.

31. The method of claim 26, wherein the polypeptide is linked to a solid phase.

32. The method of claim 26, wherein the polypeptide is expressed in a cell and the cell is contacted with the compound.

33. The method of claim 32, wherein the G-protein coupled receptor activity of the polypeptide is determined by measuring changes in intracellular cAMP, IP3, or Ca^{2+}.

34. The method of claim 26 , wherein the method comprises selecting a compound that binds to the polypeptide.

SEQUENCE LISTING

DRG1 (X2HTBKQ09Y7)

Nucleotide sequence (SEQ ID NO.: 1)

ATGGATCCAACCACCCCGGCCTGGGGAACAGAAAGTACAACAGTGAATGGAAATGACCAAGCCCTTCTTCTGCTTTGTGG
CAAGGAGACCCTGATCCCGGTCTTCCTGATCCTTTTCATTGCCCTGGTCGGGCTGGTAGGAAACGGGTTTGTGCTCTGGC
TCCTGGGCTTCCGCATGCGCAGGAACGCCTTCTCTGTCTACGTCCTCAGCGTGGCCGGGGCCGACTTCCTCTTCCTCTGC
TTCCAGATTATAAATTGCCTGGTGTACCTCAGTAACTTCTTCTGTTCCATCTCCATCAATTTCCCTAGCTTCTTCACCAC
TGTGATGACCTGTGCCTACCTTGCAGGCCTGAGCATGCTGAGCACCGTCAGCACCGAGCGCTGCCTGTCCGTCCTGTGGC
CCATCTGGTATCGCTGCCGCCGCCCCAGACACCTGTCAGCGGTCGTGTGTGTCCTGCTCTGGGCCCTGTCCCTACTGCTG
AGCATCTTGGAAGGGAAGTTCTGTGGCTTCTTATTTAGTGATGGTGACTCTGGTTGGTGTCAGACATTTGATTTCATCAC
TGCAGCGTGGCTGATTTTTTTATTCATGGTTCTCTGTGGGTCCAGT[illegible]CCTGCTGGTCAGGATCCTCTGTGGCTCCA
GGGGTCTGCCACTGACCAGGCTGTACCTGACCATCCTGCTCACAGTGCTG[illegible]TGTTCCTCCTCTGCGGCCTGCCCTTTGGC
ATTCAGTGGTTCCTAATATTATGGATCTGGAAGGATTCTGATGTCTTAT[illegible]TGTCATATTCATCCAGTTTCAGTTGTCCT
GTCATCTCTTAACAGCAGTGCCAACCCCATCATTTACTTCTTCGTGGGCTCTTTTAGGAAGCAGTGGCGGCTGCAGCAGC
CGATCCTCAAGCTGGCTCTCCAGAGGGCTCTGCAGGACATTGCTGAGGTGGATCACAGTGAAGGATGCTTCCGTCAGGGC
ACCCCGGAGATGTCGAGAAGCAGTCTGGTGTAG

Amino acid sequence (SEQ ID NO.: 2)

MDPTTPAWGTESTTVNGNDQALLLLCGKETLIPVFLILFIALVGLVGNG[illegible]LWLLGFRMRRNAFSVYVLSLAGADFLFLC
FQIINCLVYLSNFFCSISINFPSFFTTVMTCAYLAGLSMLSTVSTE[illegible]IWYRCRRPRHLSAVVCVLLWALSLLL
SILEGKFCGFLF[illegible]GDSGWCQTFDFITAAWLIFLFMVLCGSSLALL[illegible]LPLTRLYLTILLTVLVFLLCGLPFG
IQWFLILWIWKDSDVLFCHIHPVSVVLSSLNSSANPIIYFFVGSFR[illegible]LKLALQRALQDIAEVDHSEGCFRQGTPE
MSRSSLV

DRG2 (X2HTBL3EYAW)

Nucleotide sequence (SEQ ID NO.: 3)

ATGGATCCAACC[illegible]TCCCAGTCTTCGGTACAAAACTGACACCAATCAACGG[illegible]CGTGAGGAGACTCCTTGCTACAATCAGAC
CCTGAGCTTCA[illegible]TGCTGACGTG[illegible]ATCATTTCCCTTGTCGGACTGACA[illegible]AACGCGGTAGTGCTCTGGCTCCTGGGCT
ACCGCATGCGC[illegible]AACGCTGTCT[illegible]CATCTACATCCTCAACCTGG[illegible]CTTCCTCTTCCTCAGCTTCCAGATT
ATACGTTCGC[illegible]ACGCC[illegible]CATC[illegible]TATCAGCCATCTCATC[illegible]TTCTGTGATGACCTTTCCCTACTT
TACAGGCCTG[illegible]TGCTG[illegible]CATCAGCACCGAGCGCTG[illegible]GGCCCATCTGGTACCGCTGCCGCC
GCCCCACACAC[illegible]TCAGC[illegible]TGTGTCCTGCTCTGGC[illegible]TTTAGTATGCTGGAGTGGAGGTTC
TGTGACTTCCT[illegible]TAGTG[illegible]GCTGATTCTAGTTGGTGTGAA[illegible]CCCAGTCGCGTGGCTGATTTTTTT
ATGTGTGGTTC[illegible]TGTGTTTCCAG[illegible]TGGTCCTGCTGGTCAG[illegible]TCCCGGAAGATGCCGCTGACCAGGC
TGTACGTGACCA[illegible]CCTGCTCACAGTGCTGGTCTTCCTCCTCTG[illegible]CGGCATTCTGGGGGCCCTAATTTAC
AGGATGCACCT[illegible]ATTTGG[illegible]AGTCTTATATTGTCATGTTTATC[illegible]TCCCTGTCCTCTCTAAACAGTAGTGC
CAACCCCATCA[illegible]ACTTC[illegible]TCGTGGGCTCCTTTAGGCAGCGTCA[illegible]AGAACCTGAAGCTGGTTCTCCAGAGGG
CTCTGCAGGAC[illegible]CCTGA[illegible]GTG[illegible]TAAAGGTGAAGGGCAGCTTCC[illegible]AGCCTGGAGCTGTCGGGAAGCAGATTG
GGGCCATGA

Amino acid sequence (SEQ ID NO.: 4)

MDPTVPVFGT[illegible]INGREETPCY[illegible]TLSFTVLTCIISLVGLTGN[illegible]RMRRNAVSIYILNLAAADFLFLSFQI
IRSPLRLINI[illegible]RKILV[illegible]M[illegible]FTGLSMLSAISTERC[illegible]HLSAVVCVLLWGLSLLFSMLEWRF
CDFLFSGAD[illegible]TSDFI[illegible]VAW[illegible]FLCVVLCVSSLVLLVR[illegible]TILLTVLVFLLCGLPFGILGALIY
RMHLNLEVLY[illegible]LVCMS[illegible]SL[illegible]ANPIIYFFVGSFRQRQ[illegible]DKPEVDKGEGQLPEESLELSGSRLGP

DRG4 (X54[illegible]CBY8WT)

Nucleotide sequence (SEQ ID NO.: 5)

ATGTGCCCTGG[illegible]TGAGCG[illegible]GGC[illegible]CGGAACTCTACAGCCGGGGCTTCC[illegible]ATCGAGCAGATCGCGATGCTGCCGCC
TCCGGCCGTCA[illegible]AACTACATCTT[illegible]CTGCTCCTCTGCCTGTGTG[illegible]CAACGGGCTGGTCCTCTGGTTTTTCG
GCTTCTCCAT[illegible]GAGGAAC[illegible]CTCCATCTACTTCCTGCA[illegible]CGATGTGGGCTACCTCTTCAGCAAG
GCGGTGTTC[illegible]CCTG[illegible]GGGCTTC[illegible]GGGCACG[illegible]TCCGCAGCGTGTGCCGGGTCCTGGG
GCTCTGCA[illegible]TACCGGC[illegible]GCCTCCT[illegible]CGGCCG[illegible]CTCGGTCATCTTCCCCGCCT
GGTACTGG[illegible]GCGGCCCAA[illegible]CTGTCGG[illegible]CGTGGTG[illegible]TGTCCCTCCTGGTCACCTGC
CTGCACAAC[illegible]TGCGTGT[illegible]GGCCGCC[illegible]GGCCCC[illegible]GCA[illegible]TGGACATCTTCCTGGGCAT
CCTCCTGTT[illegible]CTGCT[illegible]TCATGGT[illegible]CTGCCCTC[illegible]TTGCACGTGGAGTGCCGGGCCCGAC
GGCGCCAGCC[illegible]GCCAAG[illegible]CCACGTCATCCTGGCCA[illegible]CTGGTGTCCTCCATCTACTTAGGG
ATCGACTGG[illegible]TTCTG[illegible]GT[illegible]CCAGATCCCGGCCCCCT[illegible]TCACTGACCTGTGCATCTGCATCAA

CAGCAGCGCCAAGCCCATCGTCTACTTCCTGGCCGGGAGGGACAAGTCGC[illegible]GCGGCTGTGGGAGCCGCTCAGGGTGGTCT
TCCAGCGGGCCCTGCGGGACGGCGCTGAGCTGGGGGAGGCCGGGGGCAGC[illegible]CCCAACACAGTCACCATGGAGATGCAG
TGTCCCCCGGGGAACGCCTCCTGA

Amino acid sequence (SEQ ID NO.: 6)

MCPGLSEAPEL[illegible]SRGFLTIEQIAM[illegible]PPPAVMNYIFLLLCLCGLVGNGLV[illegible]FFGFSIKRNPFSIYFLHLASADVGYLFSK
AVFSILNTGGF[illegible]TFADYIRSVCR[illegible]LGLCMFLTGVSLLPAVSAERCASV [illegible]YWRRRPKRLSAVVCALLWVLSLLVTC
LHNYFCVFLGRGA[illegible]PGAACRHMDIFLGILLFLLCCPLMVLPCLALILHV[illegible] [illegible]RQRSAKLNHVILAMVSVFLVSSIYLG
IDWFLFWVFQI[illegible]AFFPEYVTDLCICINSSAKPIVYFLAGRDKSQRLWEP[illegible] [illegible]QRALRDGAELGEAGGSTPNTVTMEMQCPP
GNAS

DRG6 (X54[illegible]RCBYJQ4)

Nucleotide sequence (SEQ ID NO.: 7)

ATGAACCAGACT[illegible]GAATAGCAGTGGGACCGTGGAGTCAGCCCTAAACTATTCCAG[illegible]GGGAGCACAGTGCACACGGCCTA
CCTGGTGCTGA[illegible]CCTGGCC[illegible]CACCTGCCTGTGCGGGATGGCAG[illegible]AC[illegible] [illegible]ATGGTGATCTGGCTGCTGGGCT
TTCGAATGC[illegible]CCCCTT[illegible]TATATCCTCAACCTGGCGGCAC[illegible] [illegible]CTTCCTCTTCAGCATGGCT
TCCACGCTCA[illegible]AAACCCCAG[illegible]TGGTCAATACCACTGACAAG[illegible] [illegible]TGAAGAGACTGATGTACTT
TGCCTACACA[illegible]CTGAGC[illegible]CGGCCATCAGCACCCAGC[illegible] [illegible]TCT[illegible]CCCTATCTGGTTCA
AGTGTCACC[illegible]CAGCACCTG[illegible]TGGGTGTGTGGCCTGCTG[illegible] [illegible]CTG[illegible]GAACGGGTTGACC
TCTTCCTTC[illegible]AAGTTCTT[illegible]TTCAATGAAGATCGGTGCTT[illegible] [illegible]CCA[illegible]CCGCCCTCATCAT
GGGGGTCT[illegible]GTGATGAC[illegible]CCAGCCTGACCCTCTTTG[illegible] [illegible]GCT[illegible]CAGCAGTGGCGGC
GGCAGCCCAC[illegible]GTTCGTGG[illegible]CTGGCCTCTGTCCTGGTGT[illegible] AT[illegible] [illegible]CCCTGCCTCTGAGCATCTAC
TGGTTTGTGCTC[illegible]ACTGGTTGAGCC[illegible]GCCCGAGATGCAGGTCCTGTG[illegible]CA[illegible] [illegible]TCACGCCTCTCCTCGTCCGT
AAGCAGCAGCC[illegible]ACCCCGTCATCTACTTCCTGGTGGGCAGCCGGAGg[illegible]AC [illegible]GCCCACCAGGTCCCTGGGGA
CTGTGCTCCA[illegible]CGCTTCG[illegible]GAGCCCGAGCTGGAAGGTGGGCACAC[illegible]CC[illegible] [illegible]TGGGCACCAATGAGATGGGG
GCTTGA

Amino acid sequence (SEQ ID NO.: 8)

MNQTLNSS[illegible]LNYSRG[illegible]TVHTAYLVLSSLAMFTCLCGMAGN[illegible]VIW[illegible]LGFPMHRNPFCIYILNLAAADLLFLFSMA
STLSLETQ[illegible]KVHEL[illegible]FAYTVGLSLLTAISTC[illegible] [illegible]HRPRHLSAWVCGLLWTLCLLMNGLT
SSFCSKF[illegible]FRVDKV[illegible]IMGVLTPVMTLSSLTLF [illegible]PTR[illegible]VVVLASVLVFLICSLPLSIY
WFVLYWLS[illegible]VLCFSLS[illegible]VSSSANPVIYFLVGSF [illegible]LQ[illegible]REEPELEGGETPTVGTNEMG
A.

DRG7 (NT[illegible]07)

Nucleotide sequence (SEQ ID NO.: 9)

ATGGATTCAA[illegible]TCCCAGTCT[illegible]GTACAGAACTGACACCAATC[illegible]ACC[illegible]GAGGAGGCGCTGACAGGAAACGCGGT
TGTGCTCTG[illegible]TGGGCTGC[illegible]TGCGCAGGAACGCTGTCTC[illegible]TC[illegible] [illegible]CCTCAACCTGGTCGCGGCCGACTTCC
TCTTCCTTA[illegible]CACATT[illegible]TCGCCGTTACGCCTCATC[illegible]AT[illegible] [illegible]ATCCCATCTCCAAAATCCTCAGTCCT
GTGATGACC[illegible]TACTT[illegible]CCCTAAGCATGCTGAGC[illegible] [illegible]GCCTGTCCATCCTGTGGCC
CATCTGGTA[illegible]CCGCC[illegible]GATACCTGTCATCAGT [illegible]GGCCCTGTCCCTGCTGCGGA
GTATCCTGG[illegible]TCTTCT[illegible]TTCCTGTTTAGTGGTG [illegible]AACGTCAGATTTCATTACA
ATCGCGTGG[illegible]TTTTT[illegible]TGGTTCTCTGTGGGTCC [illegible]GGATTCTCTGTGGATCCCG
GAAGATGCCC[illegible]CAGGC[illegible]TGACCATCCTCCTCAC [illegible]TGTGGCCTGCCCTTTGGCA
TTCAGTGGGC[illegible]TTTCCA[illegible]CACCTGGATTGGAAAG [illegible]TCTAGTTTCCATTTTCCTG
TCCGCTCTT[illegible]AGTGCC[illegible]CCATCATTTACTTCTTC [illegible]GGC[illegible]CGTCAAAATAGGCAGAACCT
GAAGCTGGT[illegible]GAGGGC[illegible]CAGGACACGCCTGAGGT[illegible] A[illegible] [illegible]GGC[illegible]GCTTCCTCAGGAAACCCTGGAGC
TGTCGGGAAC[illegible]TGGAGC[illegible]A

Amino acid sequence (SEQ ID NO.: 10)

MDSTIPVLG[illegible]INGREE[illegible]AVVLWLLGCRMRRNAVSIY[illegible]NLV[illegible]FLSGH[illegible]ICSPLRLINIRHPISKILSP
VMTFPYFIG[illegible]AISTERC[illegible]PIWYHCRRPRYLSSV[illegible] [illegible]DFLFSGANSVWCETSDFIT
IAWLVFLCV[illegible]LVLLV[illegible]RKMPLTRLYVTILLT [illegible]IHLDWKVLFCHVHLVSIFLSAL
NSSANPIIYF[illegible]RQRON[illegible]LVLQRALQDTPEVDEG[illegible] [illegible]RL[illegible]

DRG8 (NT0[illegible]282)

Nucleotide sequence (SEQ ID NO.: 11)

ATGGATCCAAC[illegible]CTCAACC[illegible]CACAGAACTGACACCAA[illegible]ACG[illegible]AC[illegible]GAGGAGACTCTTTGCTACAAGCAGAC
CTTGAGCCTC[illegible]TGCTGAC[illegible]ATCGTTTCCCTTGTCGGG[illegible]ACA[illegible] [illegible]GCA[illegible]TGTGCTCTGGCTCCTGGGCT
GCCGCATGCG[illegible]AACGCCT[illegible]CATCTACATCCTCAACTT[illegible]GC[illegible] [illegible]CTT[illegible]CTCTTCCTCAGCGGCCGCCTT

ATATATTCCCTGTTAAGCTTCATCAGTATCCCCCATACCATCTCTAAAATCCTCTATCCTGTGATGATGTTTTCCTACTT
TGCAGGCCTGAGCTTTCTGAGTGCCGTGAGCACCGAGCGCTGCCTG[illegible]CCGTCCTGTGGCCCATCTGGTACCGCTGCCACC
GCCCCACACACCTGTCAGCGGTGGTGTGTGTCCTGCTCTGGGCCC[illegible]CTGCGGAGCATCCTGGAGTGGATGTTA
TGTGGCTTCCTGTTCAGTGGTGCTGATTCTGCTTGGTGTCAAACA[illegible]ATCACAGTCGCGTGGCTGATTTTTTT
ATGTGTGGTTCTCTGTGGGTCCAGCCTGGTCCTGCTGATCAGGAT[illegible]ATCCCGGAAGATACCGCTGACCAGGC
TGTACGTGACCATCCTGCTCACAGTACTGGTCTTCCTCCTCTGTG[illegible]TTTGGCATTCAGTTTTTCCTATTTTTA
TGGATCCACGTGGACAGGGAAGTCTTATTTTGTCATGTTCATCTA[illegible]TTTCCTGTCCGCTCTTAACAGCAGTGC
CAACCCCATCATTTACTTCTTCGTGGGCTCCTTTAGGCAGCGTCA[illegible]AGAACCTGAAGCTGGTTCTCCAGAGGG
CTCTGCAGGACGCGTCTGAGGTGGATGAAGGTGGAGGGCAGCTTC[illegible]TGAGGAAATCCTGGAGCTGTCGGGAAGCAGATTGGAG
CAGTGA

Amino acid sequence (SEQ ID NO.: 12)

MDPTISTLDTELTPINGTEETLCYKQTLSLTVLTCIVSLVGLTGNAVVLWLLGCRMRRNAFSIYILNLAAADFLFLSGRL
IYSLLSFISI[illegible]TISKILYPVMM[illegible]YFAGLSFLSAVSTERCLSVL[illegible]PTHLSAVVCVLLWALSLLRSILEWML
CGFLFSGADSA[illegible]TSDFITVAWL[illegible]FLCVVLCGSSLVLLIRIL[illegible]YVTILLTVLVFLLCGLPFGIQFFLFL
WIHVDREVLFC[illegible]HLVSIFLSALN[illegible]ANPIIYFFVGSFRQRQN[illegible]LQDA[illegible]VDEGGGQLPEEILELSGSRLEQ